GMBD Core Services

Computational Proteomics

國立清華大學
生命科學系/生物資訊與結構生物研究所

Introduction to protein structure analysis- tools, databases and the web resources: ExPASy, Pfam, PDB

鄭兆勝 博士

國立清華大學

生命科學系/生物資訊與結構生物研究所

ExPASy (**Ex**pert **P**rotein **A**nalysis **Sy**stem) proteomics server

Switzerland: **http://www.expasy.org/** at Swiss Institute of Bioinformatics, Geneva
Australia: **http://au.expasy.org/** at Australian Proteome Analysis Facility, Sydney
Brazil: **http://br.expasy.org/** at Laboratório Nacional de Computação Científica, Petrópolis
Canada: **http://ca.expasy.org/** at Canadian Bioinformatics Resource, Halifax
Taiwan: **http://tw.expasy.org/** at National Health Research Institute
China: **http://cn.expasy.org/** at Peking University
Korea: **http://kr.expasy.org/** at Yonsei Proteome Research Center, Seoul

Introduction

- ExPASy是由瑞士生物資訊機構 (Swiss Institute of Bioinformatics，SIB) 所架設的伺服器。此伺服器包含了資料庫、工具&軟體、教育服務及一些相關連結。

Databases	Tools and software packages

• **UniProt Knowledgebase (Swiss-Prot and TrEMBL)** - Protein knowledgebase
• **ViralZone** - Portal to viral UniProtKB/Swiss-Prot entries
• **PROSITE** - Protein families and domains
• **SWISS-2DPAGE** - Two-dimensional polyacrylamide gel electrophoresis
• **World-2DPAGE Repository** - A public standards-compliant repository for gel-based proteomics data published in the literature
• **MIAPEGelDB** - A public repository for MIAPE Gel electrophoresis documents
• **ENZYME** - Enzyme nomenclature
• **GlycoSuiteDB** - a curated and annotated glycan database new
• **UniPathway** - Metabolic pathways
• **SWISS-MODEL Repository** - Automatically generated protein models
• Links to many other molecular biology databases
	• **Proteomics and sequence analysis tools**
	  ◦ Identification and characterization (Aldente, FindMod, Popitam, Phenyx, pI/Mw, ProtParam...)
	  ◦ DNA -> Protein
	  ◦ Similarity searches (BLAST...)
	  ◦ Pattern and profile searches (ScanProsite...)
	  ◦ Post-translational modification and topology prediction
	  ◦ Primary structure analysis
	  ◦ Secondary and tertiary structure tools (Swiss-PdbViewer...)
	  ◦ Alignment and Phylogenetic analysis
	• **Melanie / ImageMaster** - Software for 2-D PAGE analysis
	• **MSight** - Mass Spectrometry Imager
	• **Roche Applied Science's Biochemical Pathways**

Education and services	

• **The ExPASy FTP server**
• **Swiss-Shop** - automatically obtain (by email) new sequence entries relevant to your field(s) of interest
• Popular Science
  ◦ **Protein Spotlight**
  ◦ **Protéines à la «Une»**
• **e-Proxemis** - Bioinformatics Learning Portal for Proteomics
	• **Bioinformatics Core Facility for Proteomics**
	• **Proteomics Core Facility** (formerly known as SWISS-2DSERVICE)
	• **SPS' Digest** - the Swiss Proteomics Society selection of proteomics articles
	• **Master's degree in Proteomics and Bioinformatics**
	• **Swiss-Quiz**

Lists of molecular biology resources	
• **ExPASy Life Science Directory** - The ExPASy list of biomolecular servers • **BioHunt** - Search the internet for molecular biology information • **The World-2DPAGE Constellation** - Promote and publish gel-based proteomics data through the ExPASy server • **WORLD-2DPAGE List** - Links to 2-D PAGE database servers and 2-D PAGE related servers and services • **World-2DPAGE Portal** - A dynamic portal to query simultaneously World-Wide gel-based proteomics databases	
Swiss Institute of Bioinformatics (SIB) services	
• Vital-IT - The HPC Center for Life Sciences • Swiss node of EMBnet • Ashbya Genome Database • Cancer Immunome Database	• CleanEx - Gene expression database • EPD - Eukaryotic Promoter Database • GermOnLine - Knowledgebase on germ cell differentiation • smiRNAdb - a database of small RNAs • SwissRegulon - a database of regulatory sites

Database

ExPASy所提供的資料庫包含了下列幾項：

Databases

- **UniProt Knowledgebase (Swiss-Prot and TrEMBL)** - Protein knowledgebase
- **ViralZone** - Portal to viral UniProtKB/Swiss-Prot entries
- **PROSITE** - Protein families and domains
- **SWISS-2DPAGE** - Two-dimensional polyacrylamide gel electrophoresis
- **World-2DPAGE Repository** - A public standards-compliant repository for gel-based proteomics data published in the literature
- **MIAPEGelDB** - A public repository for MIAPE Gel electrophoresis documents
- **ENZYME** - Enzyme nomenclature
- **GlycoSuiteDB** - a curated and annotated glycan database new
- **UniPathway** - Metabolic pathways
- **SWISS-MODEL Repository** - Automatically generated protein models

- Links to many other molecular biology databases

UniProt Knowledgebase (Swiss-Prot and TrEMBL)

- ExPASy 本身是個資料庫的集合，主要所收集的資料是以蛋白質分子及蛋白質體組學為主的相關資料。其中 UniProt Knowledgebase (Swiss-Prot and TrEMBL)是由SIB 和EBI (歐洲分子生物學機構)所共同維護的資料庫，也就是由Swiss-Prot、TrEMBL和PIR資料庫所聯合組成的Universal Protein Knowledgebase (UniProt)聯盟。主要包含了兩部分：

  - **UniProtKB/Swiss-Prot:** 蛋白質序列資料庫，此資料庫提供高品質的注解 (annotation)，包含了蛋白質功能的描述、結構的描述、是否有後轉譯修飾(post-translational modifications)等。到2009年5月26日為止，共有468,851筆資料。

  - **UniProtKB/TrEMBL:** TrEMBL可視為對Swiss-Prot的補充，利用電腦注解的方式補充Swiss-Prot的資訊。此資料包含了將所有EMBL核酸序列轉譯成蛋白質序列，並加註於可能的功能。到2009年5月26日為止，共有7,916,844筆資料。

Notice: This page will be replaced with www.uniprot.org. Please send us your feedback!

Search for Go Clear

Swiss-Prot
Protein knowledgebase

TrEMBL
Computer-annotated supplement to Swiss-Prot

The UniProt Knowledgebase consists of:

- **UniProtKB/Swiss-Prot**; a curated protein sequence database which strives to provide a high level of annotation (such as the description of the function of a protein, its domains structure, post-translational modifications, variants, etc.), a minimal level of redundancy and high level of integration with other databases [More details / References / Linking to Swiss-Prot / User manual / Recent changes / Disclaimer].
- **UniProtKB/TrEMBL**; a computer-annotated supplement of Swiss-Prot that contains all the translations of EMBL nucleotide sequence entries not yet integrated in Swiss-Prot.

These databases are developed by the Swiss-Prot groups at SIB and at EBI.

UniProt Knowledgebase Release 15.3 consists of:
UniProtKB/Swiss-Prot Release 57.3 of 26-May-2009: 468851 entries (More statistics)
UniProtKB/TrEMBL Release 40.3 of 26-May-2009: 0 entries (More statistics)

> Swiss-Prot headlines
Rotavirus: a serial killer in UniProtKB/Swiss-Prot (Read more...)

Access to the UniProt Knowledgebase

- **New UniProt web site** new
- UniProtKB Taxonomy browser
- ViralZone - Portal to viral UniProtKB/Swiss-Prot entries new
- **BLAST** similarity search

- Retrieve a list of UniProtKB entries
- Randomly retrieve a UniProtKB entry
- UniProtKB Sequence/Annotation Version Database
- Swiss-Prot ID tracker

http://ca.expasy.org/sprot/

Prosite database

- PROSITE: Database of protein domains, families and functional sites.
- PROSITE 是由瑞士生物資訊研究所負責維護的資料庫，利用人工由文獻中蒐集資料整理而成。
- PROSITE 收集了在生物學上有顯著意義的蛋白質位點(site)和特定的序列模式(pattern)，並能根據這些位點(site)和模式(pattern)的資訊，利用電腦來進行快速鑑別工作，來鑑定一個未知功能的蛋白質序列應該屬於哪一個蛋白質家族以及可能的生物功能為何。

Prosite database

- **PROSITE patterns**
  - 一些具有生物重要性的胺基酸模式(**patterns**)。此模式可用 **regular expressions** 來表示。(e.g. [AC]-x-V-x(4)-{ED})。
- **PROSITE profiles**
  - 有些蛋白質家族 (protein families)、生物功能或結構重要的區域(domains) **無法利用被 PROSITE patterns 所歸類出來**，是因為這些蛋白質的序列是非常的 divergence，但可以利用 weight matrices (即 **profiles**) 的技術可以解決此一問題。
  - A profile is a table of position-specific amino acid weights and gap costs.
- Release 20.49, of 26-May-2009 (1555 documentation entries, 1308 patterns, 855 profiles and 861 ProRule)

ExPASy Home page | Site Map | Search ExPASy | Contact us | Swiss-Prot | ENZYME

Search PROSITE for [ ] Go Clear

Home ScanProsite ProRule Documents Downloads Links Funding

Database of protein domains, families and functional sites

PROSITE consists of documentation entries describing protein domains, families and functional sites as well as associated patterns and profiles to identify them [More details / References / Disclaimer / Commercial users].
PROSITE is complemented by **ProRule**, a collection of rules based on profiles and patterns, which increases the discriminatory power of profiles and patterns by providing additional information about functionally and/or structurally critical amino acids [More details].

Release 20.49, of 26-May-2009 (1555 documentation entries, 1308 patterns, 855 profiles and 861 ProRule)

PROSITE access

PS00747 e.g: PDOC00022, PS50089, SH3, zinc finger

Search

add wildcard '*'

輸入AC或ID或pattern name
Ex: hydrogenase or PS00747

Browse:

- by documentation entry
- by ProRule description
- by taxonomic scope
- by number of positive hit

E tools

Scan a sequence

(Output includes graphical view and feature detection)

Enter your sequence or a UniProtKB (Swiss-Prot or TrEMBL) ID or AC [help]:

Scan Clear

exclude patterns with a high probability of occurrence

- **ScanProsite** - advanced scan
- **PRATT** - allows to interactively generate conserved patterns from a series of unaligned proteins.
- **MyDomains - Image Creator** new - allows to generate custom domain figures.

http://ca.expasy.org/prosite/

Enzyme nomenclature database

- ENZYME是一個提供酵素命名相關搜尋的資料庫，而依據的酵素命名方法主要是由International Union of Biochemistry and Molecular Biology (IUBMB)提供，我們用此方法可以得到此酵素的EC (Enzyme Commission) number。

系統命名法：

1965 年命名系統化，把所有酵素依催化反應分成六大類，以四組數字名之 (IUBMB 系統)；例如 histidine carboxylase 為 EC 4.1.1.22：

Main Class：	4	Lyases	分裂 C-C, C-O, C-N 鍵
Subclass：	4.1	C-C lyase	分裂 C-C 鍵
Sub-subclass：	4.1.1	Carboxylase	分裂 C-COO 鍵
序列號碼：	22	第 22 個 4.1.1	分裂組胺酸的 C-COO 鍵

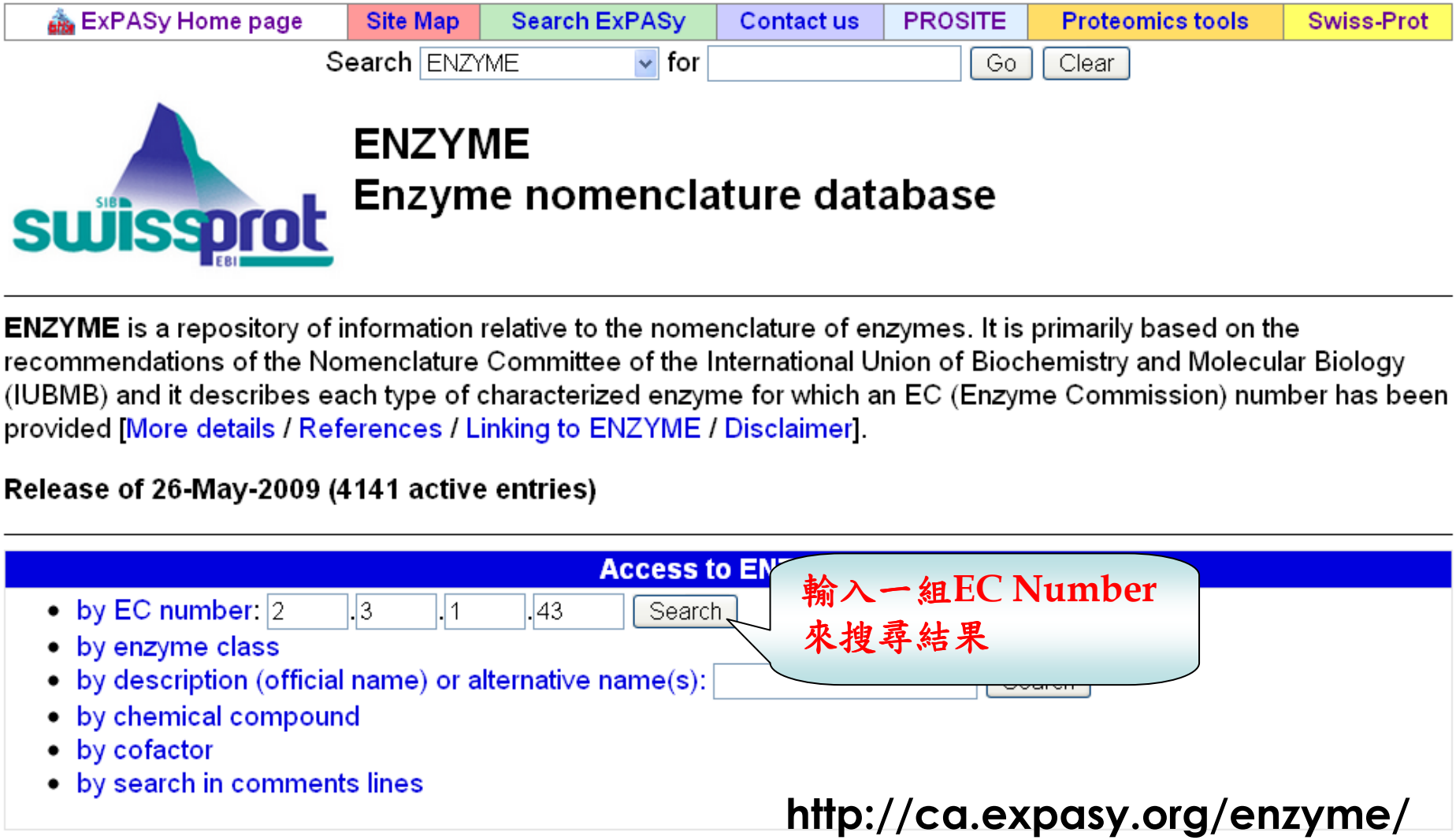
IUBMB 系統所分的六個 Main Classes：

EC1	Oxidoreductase	氧化還原酶	電子或質子轉移
EC2	Transferase	轉移酶	官能基團的轉移
EC3	Hydrolase	水解酶	加水或脫水分子
EC4	Lyase	裂解酶	共價鍵生成或裂解
EC5	Isomerase	異構酶	同一分子內基團之轉移
EC6	Ligase	連接酶	消耗 ATP 生成分子間新鍵

Access to ENZYME

- 直接輸入此酵素的**EC number**搜尋結果
- 用**Enzyme class**直接以酵素的分類方式找尋
- 輸入**Official name** 或是**Alternative name**搜尋結果
- 利用此酵素反應相關的**Chemical compound**搜尋
- 利用此酵素反應相關的**Cofactor**搜尋
- 用**Comments lines**來進行搜尋
- **SRS (Sequence Retrieval System)**

Access to ENZYME



http://ca.expasy.org/enzyme/

ExPASy
Tools and software packages

ExPASy 蛋白質體工具

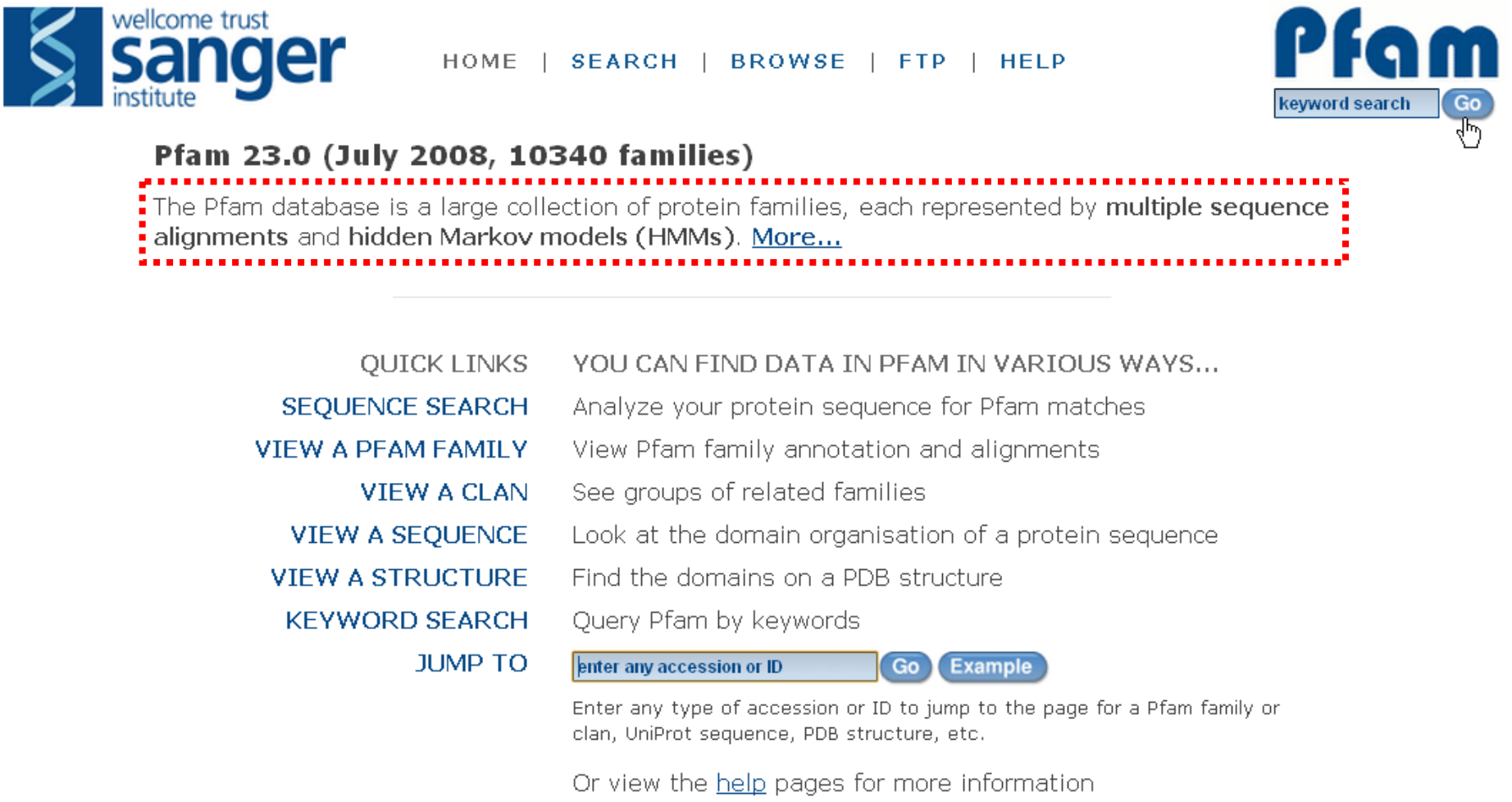
Tools and software packages

- **Proteomics and sequence analysis tools**
  - Identification and characterization (Aldente, FindMod, Popitam, Phenyx, pI/Mw, ProtParam...)
  - DNA -> Protein
  - Similarity searches (BLAST...)
  - Pattern and profile searches (ScanProsite...)
  - Post-translational modification and topology prediction
  - Primary structure analysis
  - Secondary and tertiary structure tools (Swiss-PdbViewer...)
  - Alignment and Phylogenetic analysis
- **Melanie / ImageMaster** - Software for 2-D PAGE analysis
- **MSight** - Mass Spectrometry Imager
- **Roche Applied Science's Biochemical Pathways**

Proteomics and sequence analysis tools

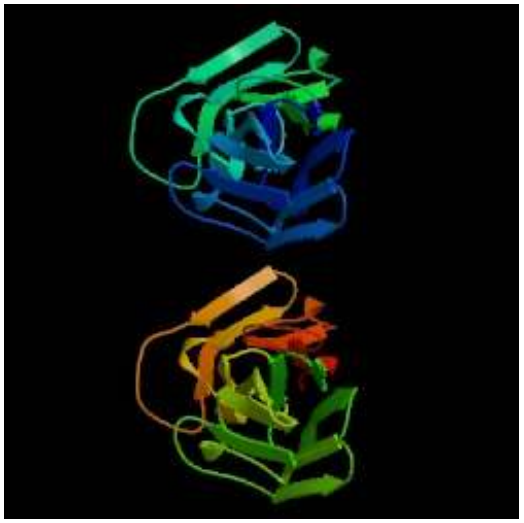
- 蛋白質身份辨識與理化特性分析 (Identification and characterization)
- 人工轉譯分析 (DNA -> Protein)
- 相似序列搜尋 (Similarity searches)
- 樣板序列搜尋與分析 (Pattern and profile searches)
- 轉譯後修飾預測 (Post-translational modification prediction)
- 拓樸特性預測分析(Topology prediction)
- 一級結構分析 (Primary structure analysis)
- 二級結構分析 (Secondary structure prediction)
- 三級結構分析 (Tertiary structure)
- 序列比對 (Sequence alignment)
- 演化樹分析 (Phylogenetic analysis)
- 生物學關鍵字分析 (Biological text analysis)

Pfam Database



http://pfam.sanger.ac.uk/

Databases for Protein Structure



Molecular Biology Database Collection

~ more than 1000 databases

- Nucleotide Sequence Databases
- RNA sequence databases
- Protein sequence databases
- Structure Databases
- Genomics Databases (non-vertebrate)
- Metabolic and Signaling Pathways
- Human and other Vertebrate Genomes
- Human Genes and Diseases
- Microarray Data and other Gene Expression Databases
- Proteomics Resources
- Other Molecular Biology Databases
- Organelle databases
- Plant databases
- Immunological databases

The Molecular Biology Database Collection: 2008 update
Nucleic Acids Research, 2008, Vol. 36, Database issue **D2-D4**

- **Structure Databases**
  - Small molecules
    - AANT - Amino Acid - Nucleotide interaction database
    - ChEBI - Chemical Entities of Biological Interest
    - ChemBank
    - ChemDB
    - CSD - Cambridge Structural Database
    - DrugBank
    - Het-PDB Navi
    - HIC-Up
    - Klotho
    - LIGAND
    - PDB-Ligand
    - PubChem
    - R.E.DD.B.
    - SuperDrug
    - SuperNatural
  - Carbohydrates
    - BCSDB/Glycoscience
    - CCSD - Complex Carbohydrate Structure Database (CarbBank)
    - CSS - Carbohydrate Structure Suite
    - Glycan
    - Glycoconjugate Data Bank
    - GlycoMapsDB
    - GlycoSuiteDB
    - Monosaccharide Browser
    - SWEET-DB

- **Structure Databases**
  - Nucleic acid structure
    - Greglist
    - GRSDB
    - ITS2
    - MeRNA
    - NCIR - Non-Canonical Interactions in RNA
    - NDB
    - NTDB
    - QuadBase
    - Rfam
    - RNA FRABASE
    - RNA SSTRAND
    - RNABase
    - RNAJunction
    - SARS-CoV RNA SSS
    - SCOR - Structural Classification Of RNA
    - Vir-Mir db
  - Protein structure
    - 3D-Genomics
    - 3DID - 3D interacting domains
    - ArchDB
    - ASTRAL
    - AutoPSI
    - BANMOKI
    - BioMagResBank
    - CADB - Conformational Angles DataBase of Proteins
    - CATH
    - CE
    - CoC Central
    - ColiSNP

- **Structure Databases**
  - Protein structure
    - Columba
    - CSA - Catalytic Site Atlas
    - Dali database
    - DBAli
    - Decoys-R-Us
    - DisProt - Database of Protein Disorder
    - DMAPS
    - Dockground
    - DomIns - Database of Domain Insertions
    - DSDBASE - Disulfide Database
    - DSMM - a Database of Simulated Molecular Motions
    - E-MSD - EBI-Macromolecular Structure Database
    - eF-site - Electrostatic surface of Functional site
    - EzCatDB
    - FireDB
    - FSN
    - Gene3D
    - Genomic Threading Database
    - GTOP - Genomes To Protein structures
    - HOMSTRAD - Homologous Structure Alignment Database
    - HotSprint
    - IMB Jena Image Library
    - IMGT/3Dstructure-DB
    - IMOTdb
    - MALISAM
    - LPFC
    - MegaMotifbase
    - MMDB
    - ModBase
    - MolMovDB - Database of Macromolecular Movements
    - PASS2

- **Structure Databases**
  - Protein structure
    - PDB
    - PDB-REPRDB
    - PDBselect
    - PDBsum
    - PDB_TM
    - PepConfDB
    - PFD - Protein Folding Database
    - Phospho3D
    - PIDD
    - PMDB - Protein Model Database
    - Structure Superposition Database
    - ProSAS
    - PROTCOM
    - PRTAD
    - RESID
    - S4: Structure-based Sequence Alignments of SCOP Superfamilies
    - SCOP - Structural Classification Of Proteins
    - SCOPPI
    - SitesBase
    - SNAPPI
    - SSToSS - Sequence-Structural Templates of Single-member Superfamilies
    - STINGreport
    - SUPERFAMILY
    - SURFACE
    - SWISS-MODEL Repository
    - TargetDB
    - TMBETA-GENOME
    - TOPOFIT-DB
    - TOPS - Topology Of Protein Structures

Protein Data Bank (PDB)

- http://**www.pdb.org/**
- Structure data determined by X-ray crystallography and NMR
- The data include the atom coordinate, reference, sequence, secondary structure, disulfide bond ……etc.

The number of protein structure and the last update date

MyPDB Login | A MEMBER OF THE PDB

An Information Portal to Biological Macromolecular Structures

As of **Tuesday Jun 02, 2009** there are 57944 Structures | **PDB Statistics**

WHAT'S NEW | CONTACT | FEEDBACK | HELP | PRINT | PDB ID or keyword | Adv. Search

PDB Statistics
57944 Structures
Last Update: **Jun 02, 2009**

Home | Search

- Home
- Getting Started
- Structural Genomics
- BioSync
- Electron Microscopy
- Download Files
- Deposit and Validate
- Dictionaries & File Formats
- Software Tools
- General Education
- General Information
- Acknowledgements
- Frequently Asked Questions

A Resource for Studying Biological Macromolecules

The PDB archive contains information about experimentally-determined structures of proteins, nucleic acids, and complex assemblies. As a member of the **wwPDB**, the RCSB PDB curates and annotates PDB data according to agreed upon standards.

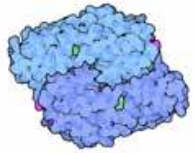
The RCSB PDB also provides a variety of tools and resources. Users can perform simple and advanced searches based on annotations relating to sequence, structure and function. These molecules are visualized, downloaded, and analyzed by users who range from students to specialized scientists.

Molecule of the Month: Vaults

Our cells are filled with compartments, each performing a specific function. Some of these compartments, such as mitochondria and lysozomes, are very large and enclose many different molecular machines. Other intracellular compartments are smaller, such as the transport vesicles that shuttle proteins from site to site inside the cell. Most of these compartments, including mitochondria, lysozomes and transport vesicles, are surrounded by membranes. However, in special cases, cells build smaller compartments surrounded by a protein shell. In our own cells, vaults are a spectacular example of these protein-enclosed compartments.

Read more ... | **Previous Features**

PSI Featured Molecule: Hda and DNA Replication



When cells divide, they need to ensure that each daughter cell gets one copy of each chromosome. Bacteria contain one big circle of DNA, so they start replication in one place, then copy the DNA both ways around until it finishes on the other side. PSI Researchers have solved the first atomic structure at how bacteria use the Hda protein to initiate replication at this origin only once for each generation of the cell.

Read more from PSI SGKB | **Previous Features**

New user? Try the browser **compatibility check**, information on **Getting Started**, and see **What's New** on the site.

- **Newsletter**
- **Discussion Forum**
- **Job Listings**

02-June-2009

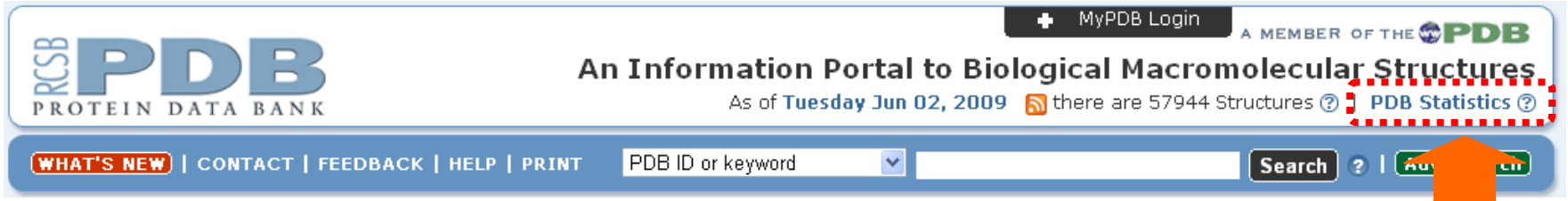
Literature View: Looking at Structures in PubMedCentral

The recently-released Literature View aims to provide a broad look at how a given structure has been analyzed and presented in open access publications, that is, where the full text of the article is available without copyright restriction. The overall intent is to make the PDB user more aware of publications associated with the structure under study.

More >>

Data Snapshots

Time-stamped yearly snapshots of the PDB archive are available via FTP at: **ftp://snapshots.wwpdb.org** The snapshots provide readily identifiable data sets for research on the PDB archive.



PDB Current Holdings Breakdown

		Proteins	Nucleic Acids	Protein/NA Complexes	Other	Total
Exp. Method	X-ray	46383	1147	2141	17	49688
	NMR	6864	856	146	6	7872
	Electron Microscopy	168	16	59	0	243
	Hybrid	13	1	1	1	16
	Other	108	4	4	9	125
	Total	53536	2024	2351	33	57944

(Click on any number to retrieve the results from that category.)

Please note that theoretical models have been removed, effective July 02, 2002, as per **PDB policy**.

38858 structures in the PDB have a structure factor file.
5156 structures in the PDB have an NMR restraint file.

- As Superfamilies Defined By SCOP
- As Superfamilies Defined By CATH

Use **Search Unreleased** to search and view entries that are currently being processed or are awaiting release.

Statistics are for experimentally-determined structures.

RCSB PDB
PROTEIN DATA BANK

An Information Portal

As of Tuesday Jun (

WHAT'S NEW | CONTACT | FEEDBACK | HELP | PRINT PDB ID or keyword

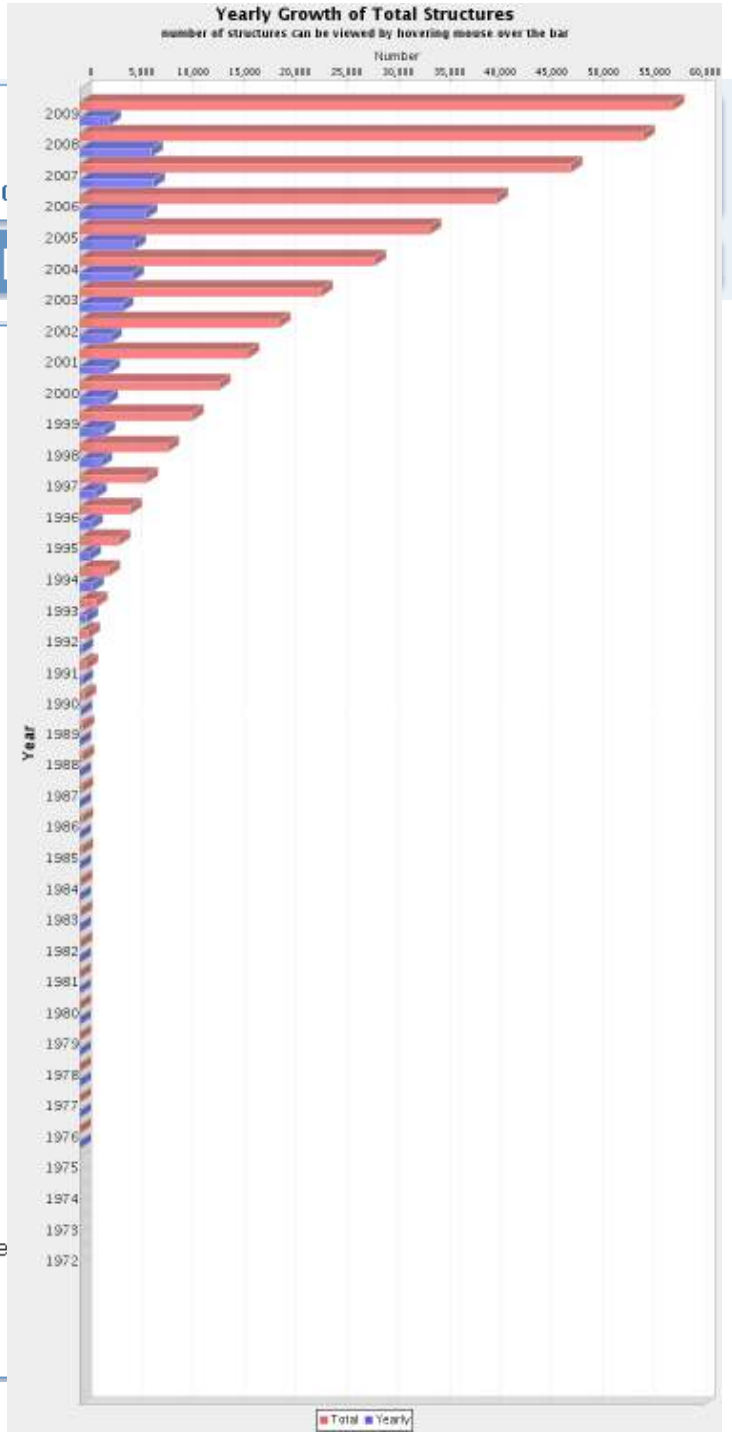
PDB Statistics

- Content Distribution
  - Summary Table of Released Entries
  - Status of Unreleased Entries
  - Proteins solved by multiple experimental methods
  - Redundancy based on sequence similarity
  - By Resolution
  - By Space Group
  - By Natural Source Organism
  - By Gene Source Scientific Organism
  - By Top 100 Journals
  - By Structural Genomics Centers
  - By Structure Molecular Weight
  - By Enzyme Classification

- Content Growth
  - Growth of Released Structures Per Year
  - Growth of Released Structures Per Year By Experimental Method
    - X-ray
    - NMR
    - Electron Microscopy
  - Growth of Released Structures Per Year By Molecular Type
    - Protein Only
    - DNA Only
    - RNA Only
    - Protein Nucleic Acid Complexes
  - Growth Of Unique Protein Classifications Per Year
    - As Folds Defined By SCOP
    - As Topologies Defined By CATH
    - As Superfamilies Defined By SCOP
    - As Superfamilies Defined By CATH

Use Search Unreleased to search and view entries that are curre release.

Statistics are for experimentally-determined structures.



Tutorial About This Site

RCSB PDB
PROTEIN DATA BANK

WHAT'S NEW | CONTACT | FEEDBACK | HELP | PRINT

Home Search

- Home
- 1 Getting Started
- Structural Genomics
- BioSync
- Electron Microscopy
- Download Files
- Deposit and Validate
- Dictionaries & File Formats
- Software Tools
- General Education
- General Information
- Acknowledgements
- Frequently Asked Questions

Getting Started

The RCSB PDB website provides

Quick Start

Navigation: To navigate through
General Education section of t

2 A narrated tutorial illustra
also provides a good introductio

Searching: To search for struct
more detailed search options, su

Browsers and software: **Most**
browsers. A java-enabled brows
browser and browser settings ar

Some molecular viewers may re

A Flash player is required for tut

Website Organiza

Top Search Bar: The top frame
home page. Selecting the date
help desk.

Left Menu: Clicking on an arro
related to that topic. Clicking on

Flash

http://www.pdb.org/pdb/tutorials/Main_Tutorial.swf

Search

RCSB PDB PROTEIN DATA BANK

MyPDB Login

A MEMBER OF THE PDB

Information Portal to Biological Macromolecular Structures

As of Tuesday Jun 02, 2009 there are 57944 Structures | PDB Statistics

WHAT'S NEW | CONTACT | FEEDBACK | HELP | PRINT

PDB ID or keyword | Search | Adv. Search

PDB ID or keyword
Author
Structural Genomics Center
Chemical Name
Chemical ID

Home | Search

- Home
- Getting Started
- Structural Genomics
- BioSync
- Electron Microscopy
- Download Files
- Deposit and Validate
- Dictionaries & File Formats
- Software Tools
- General Education
- General Information
- Acknowledgements
- Frequently Asked Questions

A Resource for Stud... Macromolecules

The PDB archive contains information abou... acids, and complex assemblies. As a mem... PDB data according to agreed upon stand...

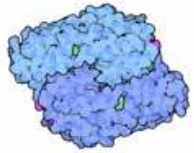
The RCSB PDB also provides a variety... advanced searches based on annotati... molecules are visualized, downloaded,... specialized scientists.

Molecule of the Month: Vaults

Our cells are filled with compartments, each performing a specific function. Some of these compartments, such as mitochondria and lysozomes, are very large and enclose many different molecular machines. Other intracellular compartments are smaller, such as the transport vesicles that shuttle proteins from site to site inside the cell. Most of these compartments, including mitochondria, lysozomes and transport vesicles, are surrounded by membranes. However, in special cases, cells build smaller compartments surrounded by a protein shell. In our own cells, vaults are a spectacular example of these protein-enclosed compartments.

Read more ... | Previous Features

PSI Featured Molecule: Hda and DNA Replication



When cells divide, they need to ensure that each daughter cell gets one copy of each chromosome. Bacteria contain one big circle of DNA, so they start replication in one place, then copy the DNA both ways around until it finishes on the other side. PSI Researchers have solved the first atomic structure at how bacteria use the Hda protein to initiate replication at this origin only once for each generation of the cell.

Read more from PSI SGKB | Previous Features

New user? Try the browser **compatibility check**, information on **Getting Started**, and see **What's New** on the site.

...News

...Forum

...w: Looking at ...al

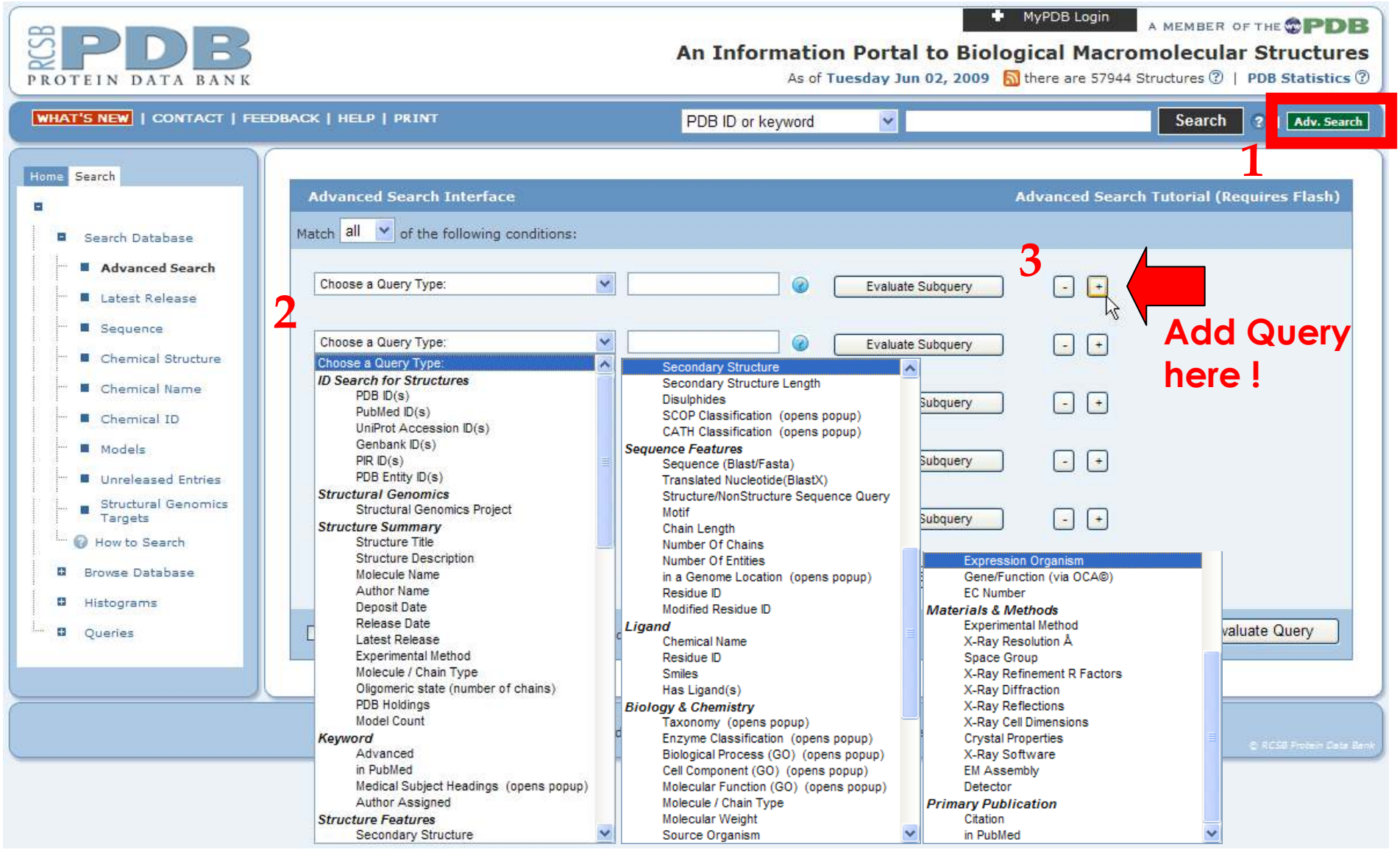
The recently-released Literature View aims to provide a broad look at how a given structure has been analyzed and presented in open access publications, that is, where the full text of the article is available without copyright restriction. The overall intent is to make the PDB user more aware of publications associated with the structure under study.

More >>

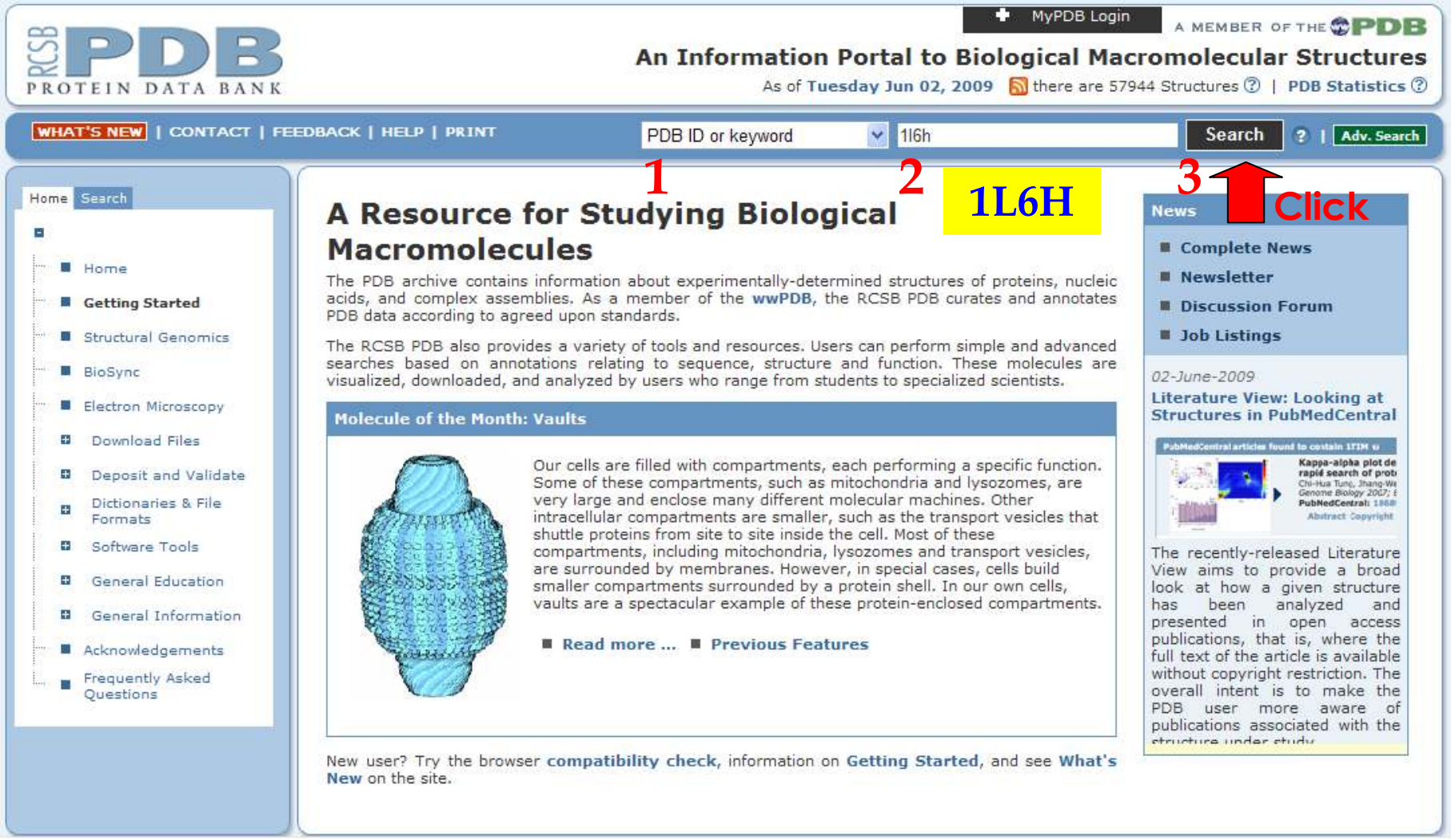
Data Snapshots

Time-stamped yearly snapshots of the PDB archive are available via FTP at: **ftp://snapshots.wwpdb.org** The snapshots provide readily identifiable data sets for research on the PDB archive.

Advanced Search Interface

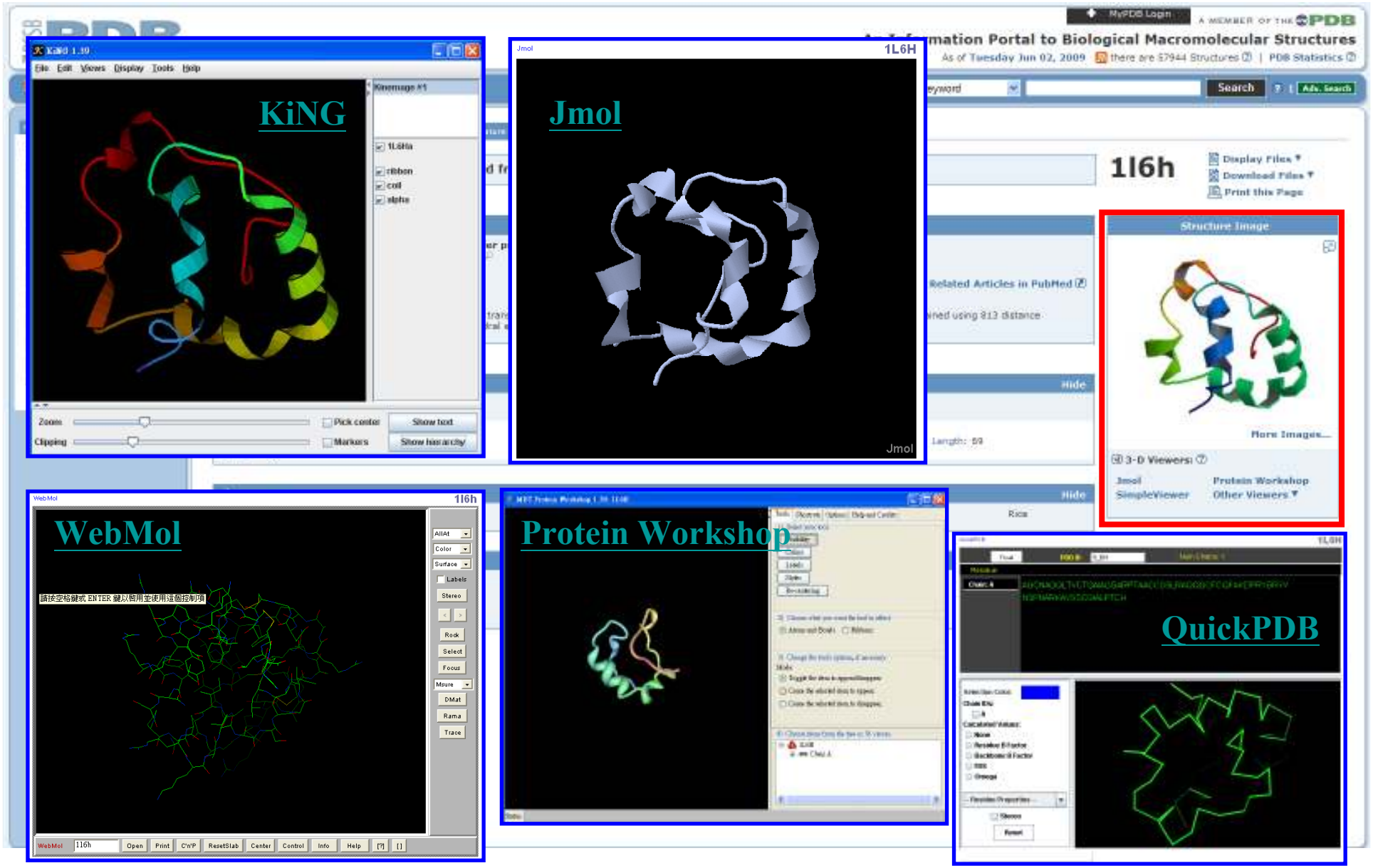


Search PDB ID: 1L6H

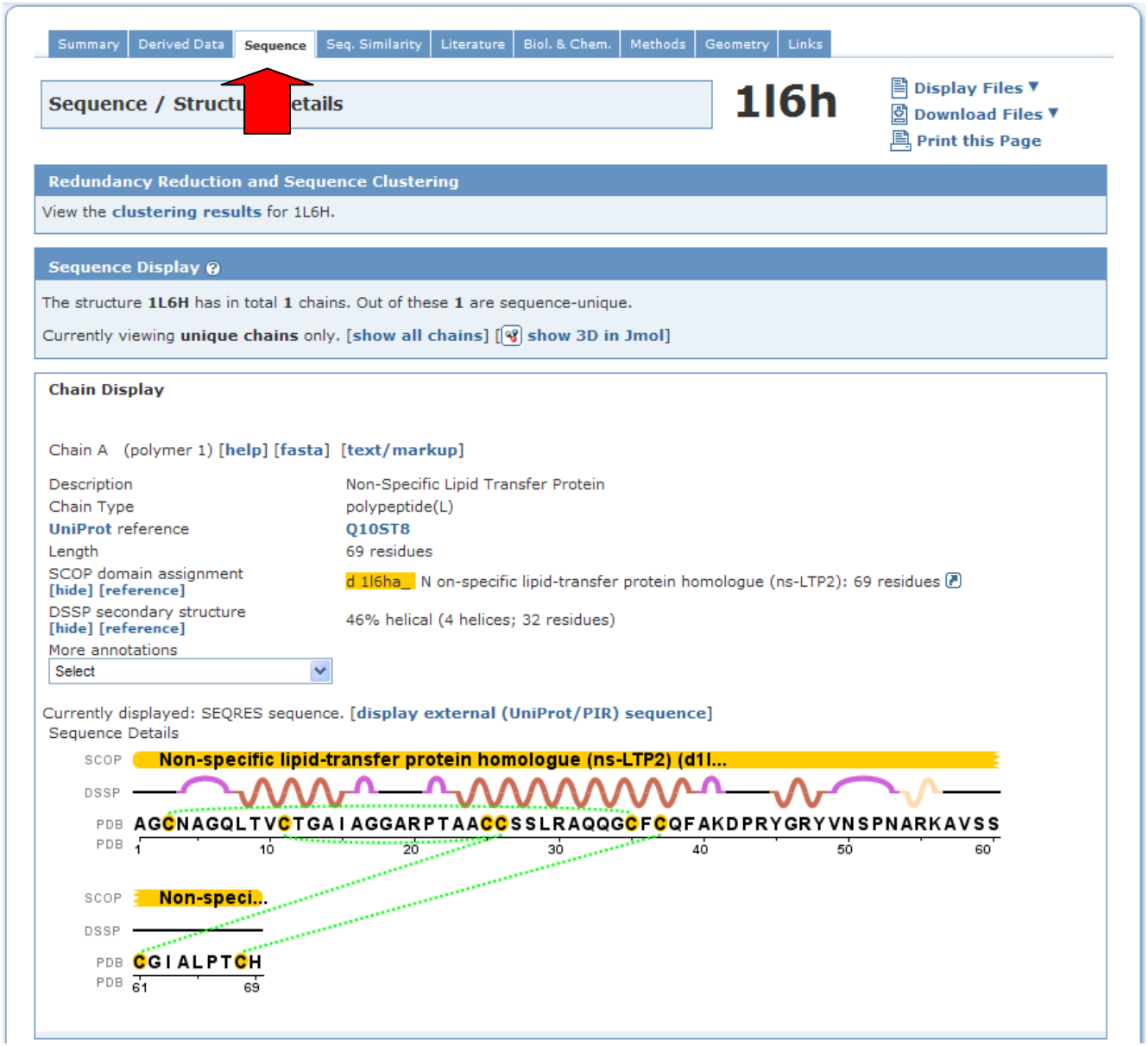


Structure Summary : 1L6H

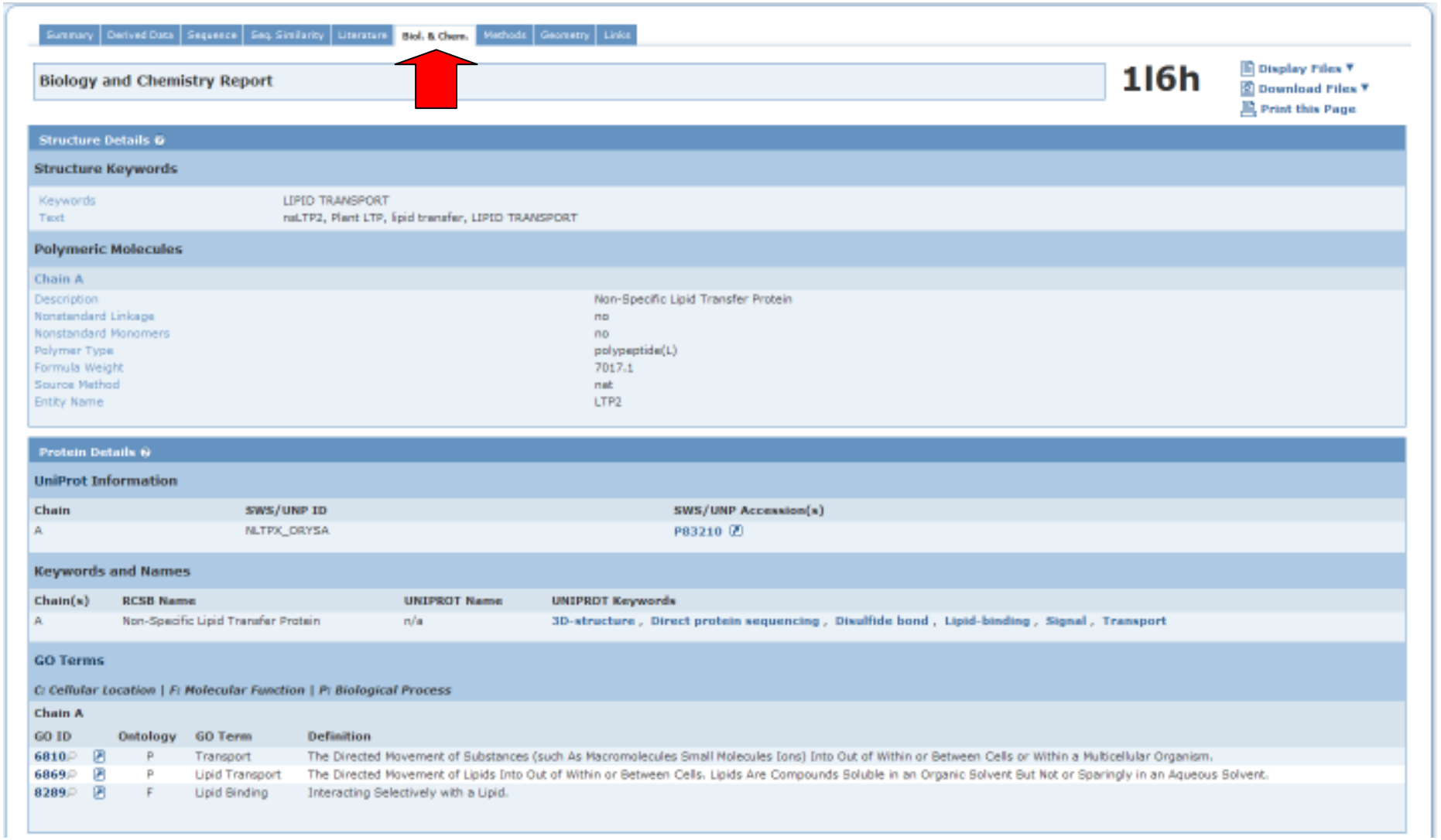
View Structure : 1L6H



Sequence / Structure Details



Biology and Chemistry Report



Summary | Derived Data | Sequence | Seq. Similarity | Literature | Biol. & Chem. | Methods | Geometry | Links

Biology and Chemistry Report

1I6h

Display Files ▾
Download Files ▾
Print this Page

Structure Details

Structure Keywords

Keywords	LIPID TRANSPORT
Text	nsLTP2, Plant LTP, lipid transfer, LIPID TRANSPORT

Polymeric Molecules

Chain A

Description	Non-Specific Lipid Transfer Protein
Nonstandard Linkage	no
Nonstandard Monomers	no
Polymer Type	polypeptide(L)
Formula Weight	7017.1
Source Method	nat
Entity Name	LTP2

Protein Details

UniProt Information

Chain	SWS/UNP ID	SWS/UNP Accession(s)
A	NLTPX_ORYSA	P83210

Keywords and Names

Chain(s)	RCSB Name	UNIPROT Name	UNIPROT Keywords
A	Non-Specific Lipid Transfer Protein	n/a	3D-structure , Direct protein sequencing , Disulfide bond , Lipid-binding , Signal , Transport

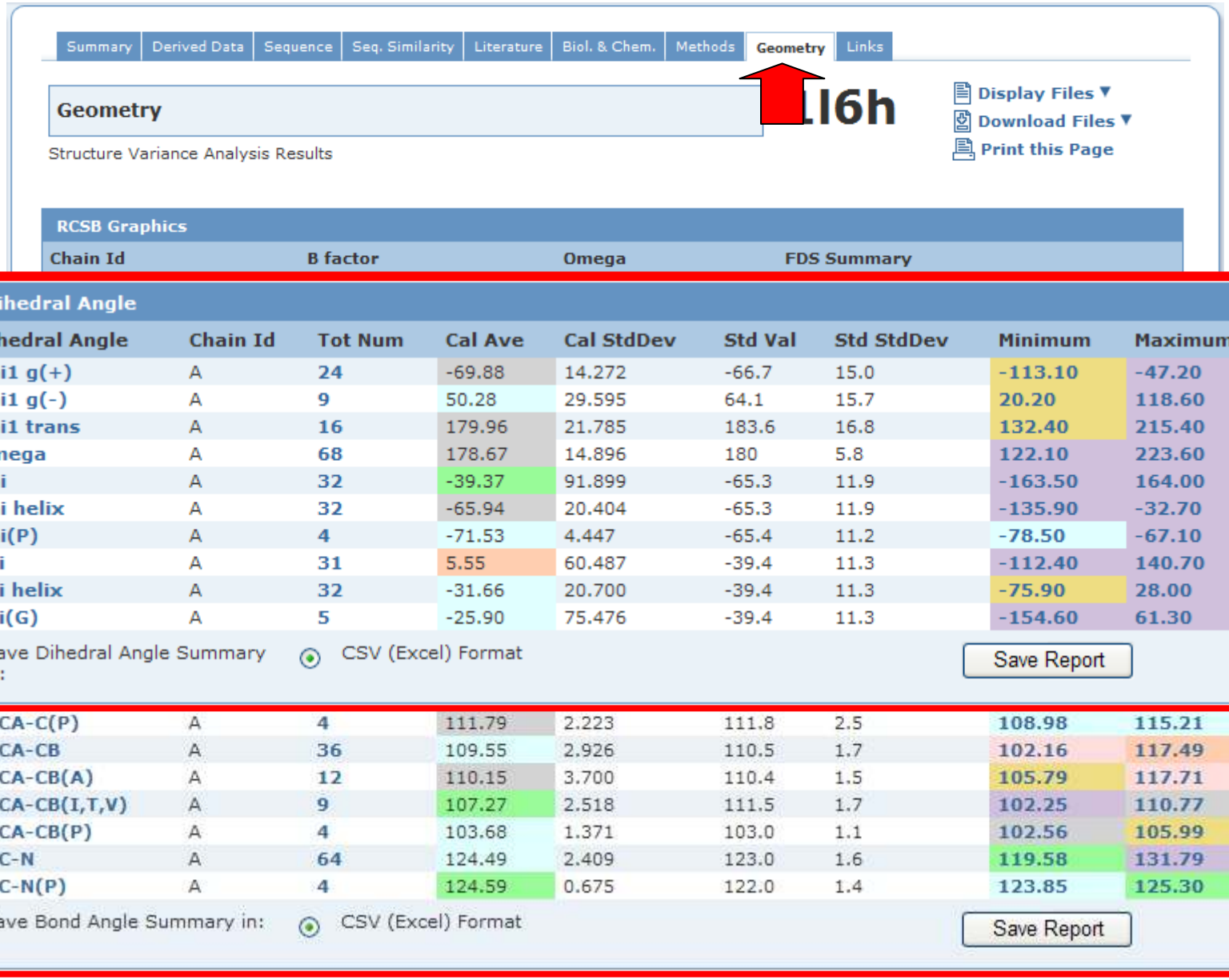
GO Terms

C: Cellular Location | F: Molecular Function | P: Biological Process

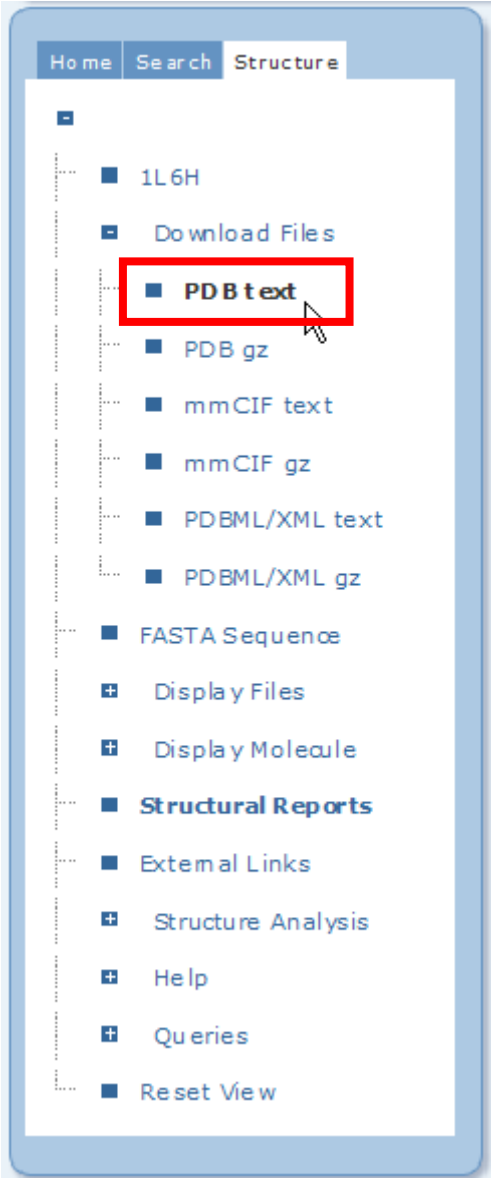
Chain A

GO ID	Ontology	GO Term	Definition
6810	P	Transport	The Directed Movement of Substances (such As Macromolecules Small Molecules Ions) Into Out of Within or Between Cells or Within a Multicellular Organism.
6869	P	Lipid Transport	The Directed Movement of Lipids Into Out of Within or Between Cells. Lipids Are Compounds Soluble in an Organic Solvent But Not or Sparingly in an Aqueous Solvent.
8289	F	Lipid Binding	Interacting Selectively with a Lipid.

Geometry

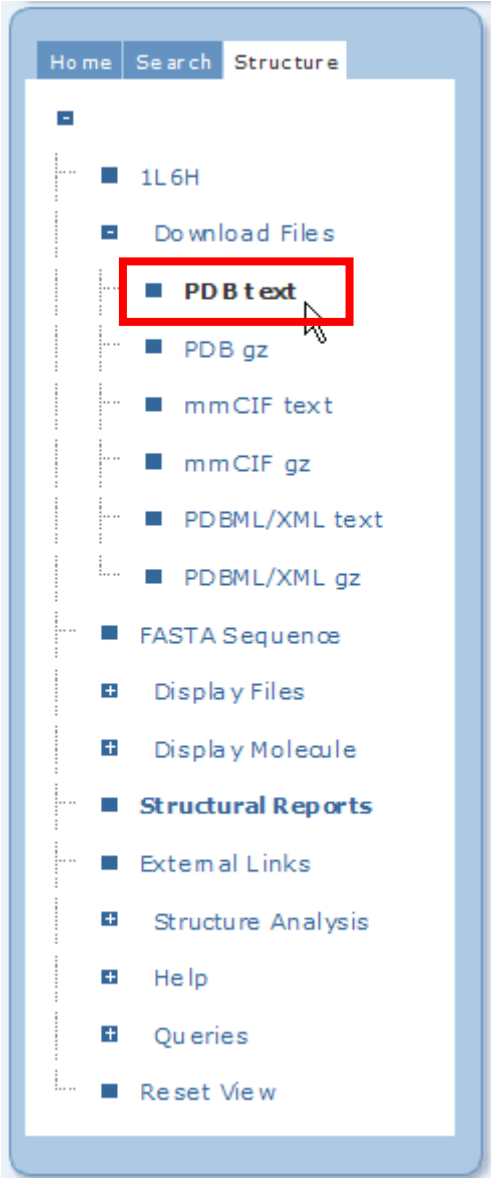

Summary
Derived Data
Sequence
Seq. Similarity
Literature
Biol. & Chem.
Methods
Geometry
Links
Geometry
Structure Variance Analysis Results
Display Files
Download Files
Print this Page
RCSB Graphics
Chain Id
B factor
Omega
FDS Summary
Dihedral Angle
Dihedral Angle | Chain Id | Tot Num | Cal Ave | Cal StdDev | Std Val | Std StdDev | Minimum | Maximum
Chi1 g(+) | A | 24 | -69.88 | 14.272 | -66.7 | 15.0 | -113.10 | -47.20
Chi1 g(-) | A | 9 | 50.28 | 29.595 | 64.1 | 15.7 | 20.20 | 118.60
Chi1 trans | A | 16 | 179.96 | 21.785 | 183.6 | 16.8 | 132.40 | 215.40
Omega | A | 68 | 178.67 | 14.896 | 180 | 5.8 | 122.10 | 223.60
Phi | A | 32 | -39.37 | 91.899 | -65.3 | 11.9 | -163.50 | 164.00
Phi helix | A | 32 | -65.94 | 20.404 | -65.3 | 11.9 | -135.90 | -32.70
Phi(P) | A | 4 | -71.53 | 4.447 | -65.4 | 11.2 | -78.50 | -67.10
Psi | A | 31 | 5.55 | 60.487 | -39.4 | 11.3 | -112.40 | 140.70
Psi helix | A | 32 | -31.66 | 20.700 | -39.4 | 11.3 | -75.90 | 28.00
Psi(G) | A | 5 | -25.90 | 75.476 | -39.4 | 11.3 | -154.60 | 61.30
Save Dihedral Angle Summary in:
CSV (Excel) Format
Save Report
N-CA-C(P) | A | 4 | 111.79 | 2.223 | 111.8 | 2.5 | 108.98 | 115.21
N-CA-CB | A | 36 | 109.55 | 2.926 | 110.5 | 1.7 | 102.16 | 117.49
N-CA-CB(A) | A | 12 | 110.15 | 3.700 | 110.4 | 1.5 | 105.79 | 117.71
N-CA-CB(I,T,V) | A | 9 | 107.27 | 2.518 | 111.5 | 1.7 | 102.25 | 110.77
N-CA-CB(P) | A | 4 | 103.68 | 1.371 | 103.0 | 1.1 | 102.56 | 105.99
O-C-N | A | 64 | 124.49 | 2.409 | 123.0 | 1.6 | 119.58 | 131.79
O-C-N(P) | A | 4 | 124.59 | 0.675 | 122.0 | 1.4 | 123.85 | 125.30
Save Bond Angle Summary in:
CSV (Excel) Format
Save Report

Download/Display File




```
HEADER    LIPID TRANSPORT                 11-MAR-02   1L6H
TITLE     SOLUTION STRUCTURE OF PLANT NSLTP2 PURIFIED FROM RICE
TITLE    2 (ORYZA SATIVA)
COMPND    MOL_ID: 1;
COMPND   2 MOLECULE: NON-SPECIFIC LIPID TRANSFER PROTEIN;
COMPND   3 CHAIN: A;
COMPND   4 SYNONYM: LTP2
SOURCE    MOL_ID: 1;
SOURCE   2 ORGANISM_SCIENTIFIC: ORYZA SATIVA;
SOURCE   3 ORGANISM_COMMON: RICE;
SOURCE   4 ORGANISM_TAXID: 4530
KEYWDS    NSLTP2, PLANT LTP, LIPID TRANSFER, LIPID TRANSPORT
EXPDTA    SOLUTION NMR
MDLTYP    MINIMIZED AVERAGE
AUTHOR    D.SAMUEL,P.-C.LYU
REVDAT   4   24-FEB-09 1L6H    1       VERSN
REVDAT   3   01-APR-03 1L6H    1       JRNL
REVDAT   2   04-FEB-03 1L6H    1       REMARK
REVDAT   1   02-OCT-02 1L6H    0
JRNL      AUTH   D.SAMUEL,Y.J.LIU,C.S.CHENG,P.C.LYU
JRNL      TITL   SOLUTION STRUCTURE OF PLANT NONSPECIFIC LIPID
JRNL      TITL 2 TRANSFER PROTEIN-2 FROM RICE (ORYZA SATIVA).
JRNL      REF    J.BIOL.CHEM.                V. 277 35267 2002
JRNL      REFN                   ISSN 0021-9258
JRNL      PMID   12011089
JRNL      DOI    10.1074/JBC.M203113200
REMARK   1
REMARK   1 REFERENCE 1
REMARK   1  AUTH   Y.-J.LIU,D.SAMUEL,C.-H.LIN,P.-C.LYU
REMARK   1  TITL   PURIFICATION AND CHARACTERIZATION OF A NOVEL 7-KDA
REMARK   1  TITL 2 NON-SPECIFIC LIPID TRANSFER PROTEIN-2 FROM RICE
REMARK   1  TITL 3 (ORYZA SATIVA)
REMARK   1  REF    BIOCHEM.BIOPHYS.RES.COMMUN.   V. 294   535 2002
REMARK   1  REFN                   ISSN 0006-291X
REMARK   1  DOI    10.1016/S0006-291X(02)00509-0
REMARK   2
REMARK   2 RESOLUTION. NOT APPLICABLE.
```

Download/Display File



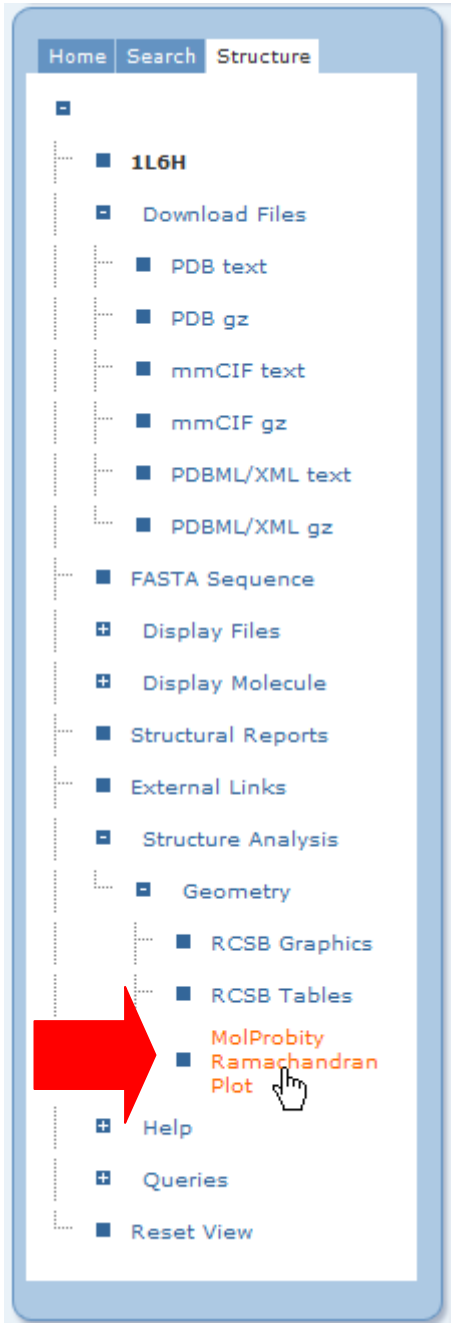

```
HELIX   1  1 GLN A   7  ALA A  16  1                              10
HELIX   2  2 THR A  22  ALA A  40  1                              19
HELIX   3  3 ARG A  44  VAL A  49  1                               6
HELIX   4  4 ASN A  53  ALA A  57  5                               5
SSBOND  1 CYS A   3  CYS A  35                          1555  1555  2.02
SSBOND  2 CYS A  11  CYS A  25                          1555  1555  2.04
SSBOND  3 CYS A  26  CYS A  61                          1555  1555  2.04
SSBOND  4 CYS A  37  CYS A  68                          1555  1555  2.05
CRYST1  1.000   1.000   1.000  90.00  90.00  90.00 P 1          1
ORIGX1      1.000000  0.000000  0.000000        0.00000
ORIGX2      0.000000  1.000000  0.000000        0.00000
ORIGX3      0.000000  0.000000  1.000000        0.00000
SCALE1      1.000000  0.000000  0.000000        0.00000
SCALE2      0.000000  1.000000  0.000000        0.00000
SCALE3      0.000000  0.000000  1.000000        0.00000
ATOM      1  N   ALA A   1      42.576 -11.286  -3.123  1.00  0.00           N
ATOM      2  CA  ALA A   1      42.743  -9.879  -3.389  1.00  0.00           C
ATOM      3  C   ALA A   1      43.818  -9.551  -2.395  1.00  0.00           C
ATOM      4  O   ALA A   1      43.756 -10.044  -1.336  1.00  0.00           O
ATOM      5  CB  ALA A   1      41.585  -8.965  -2.979  1.00  0.00           C
ATOM      6  H1  ALA A   1      41.736 -11.706  -3.571  1.00  0.00           H
ATOM      7  H2  ALA A   1      42.601 -11.344  -2.083  1.00  0.00           H
ATOM      8  H3  ALA A   1      43.464 -11.717  -3.446  1.00  0.00           H
ATOM      9  HA  ALA A   1      43.061  -9.768  -4.410  1.00  0.00           H
ATOM     10  HB1 ALA A   1      41.914  -7.920  -3.005  1.00  0.00           H
ATOM     11  HB2 ALA A   1      41.274  -9.133  -1.949  1.00  0.00           H
ATOM     12  HB3 ALA A   1      40.721  -9.055  -3.633  1.00  0.00           H
ATOM     13  N   GLY A   2      44.716  -8.681  -2.942  1.00  0.00           N
ATOM     14  CA  GLY A   2      45.682  -7.887  -2.232  1.00  0.00           C
ATOM     15  C   GLY A   2      46.351  -8.666  -1.139  1.00  0.00           C
ATOM     16  O   GLY A   2      46.368  -9.866  -1.120  1.00  0.00           O
ATOM     17  H   GLY A   2      44.679  -8.358  -3.861  1.00  0.00           H
ATOM     18  HA2 GLY A   2      46.434  -7.585  -2.943  1.00  0.00           H
ATOM     19  HA3 GLY A   2      45.157  -7.003  -1.881  1.00  0.00           H
ATOM     20  N   CYS A   3      46.781  -7.815  -0.201  1.00  0.00           N
```

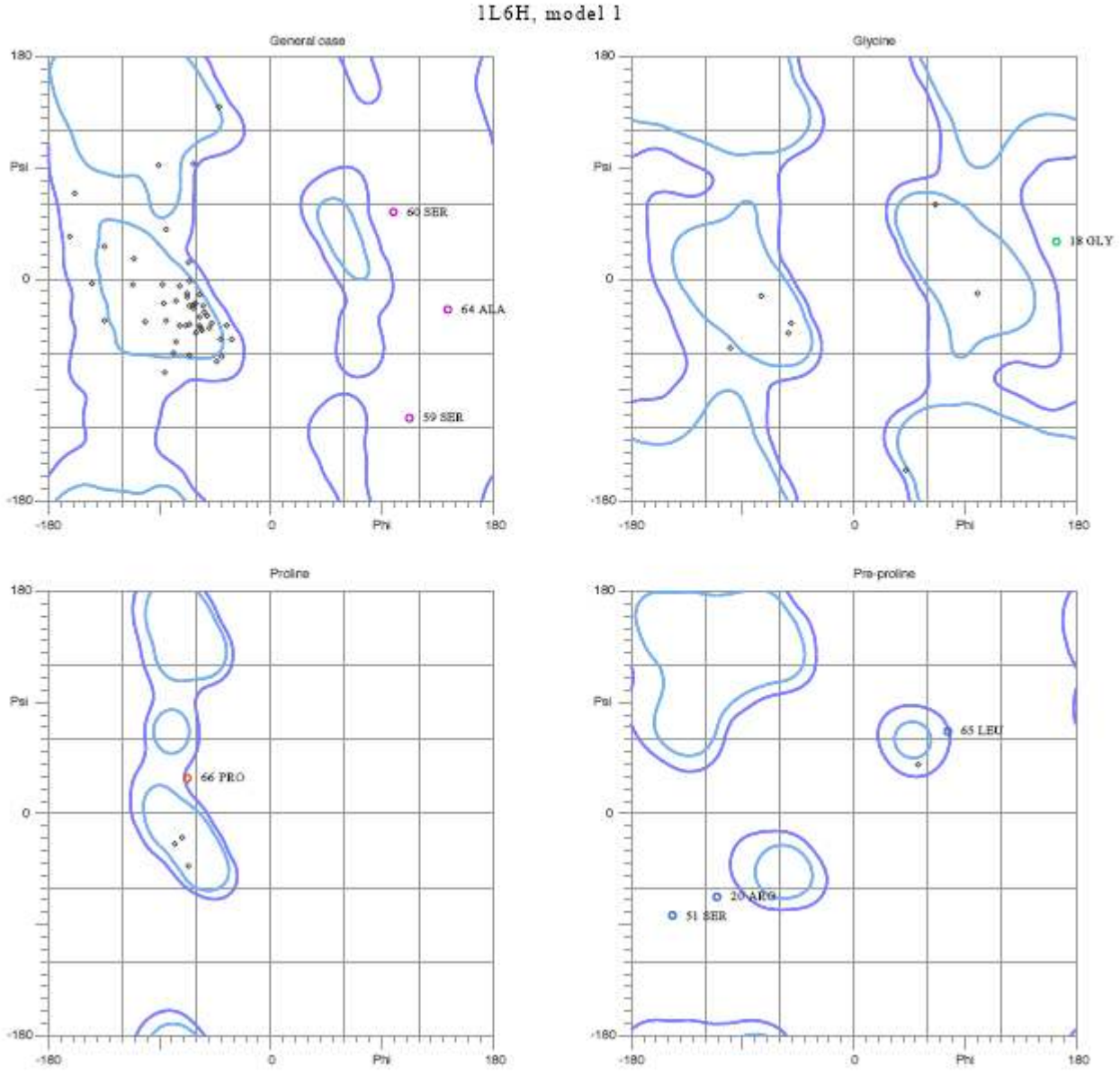
x y z

PDB File Title Section

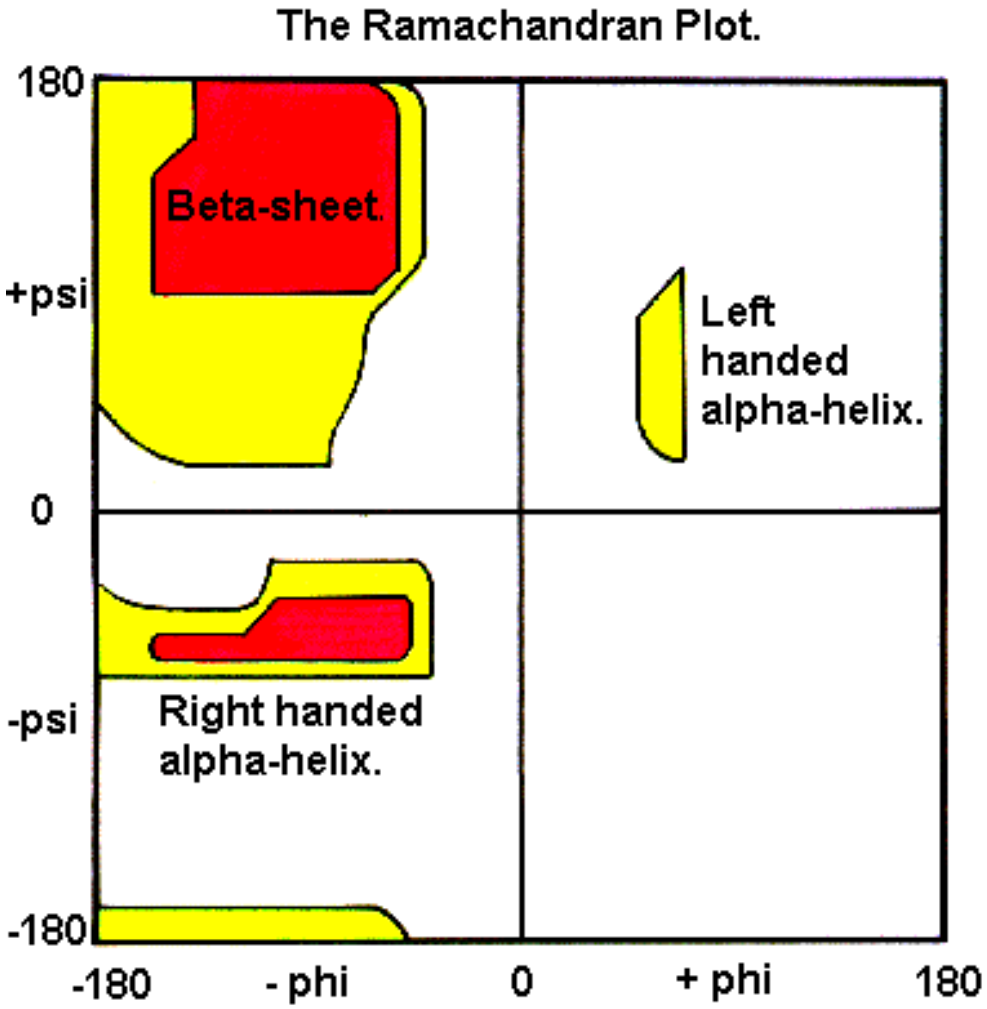
HEAD	First line of the entry, contains PDB ID code, classification, and date of deposition.	**HELIX**	Identification of helical substructures.
COMPND	Description of macromolecular contents of the entry.	**CRYST1**	Unit cell parameters, space group, and Z.
SOURCE	Biological source of macromolecules in the entry.	**ORIGXn**	Transformation from orthogonal coordinates to the submitted coordinates (n = 1, 2, or 3).
AUTHOR	List of contributors.	**SCALEn**	Transformation from orthogonal coordinates to fractional crystallographic coordinates (n = 1, 2, or 3).
REVDAT	Revision date and related information.	**MTRIXn**	Transformations expressing non-crystallographic symmetry (n = 1, 2, or 3). There may be multiple sets of these records.
JRNL	Literature citation that defines the coordinate set.	**ATOM**	Atomic coordinate records for standard groups.
REMARK	General remarks, some are structured and some are free form.	**HETATM**	Atomic coordinate records for heterogens.
SEQRES	Primary sequence of backbone residues.	**TER**	Chain terminator.
FORMUL	Chemical formula of non-standard groups.	**END**	Last record in the file.

Ramachandran Plot





Ramachandran plot



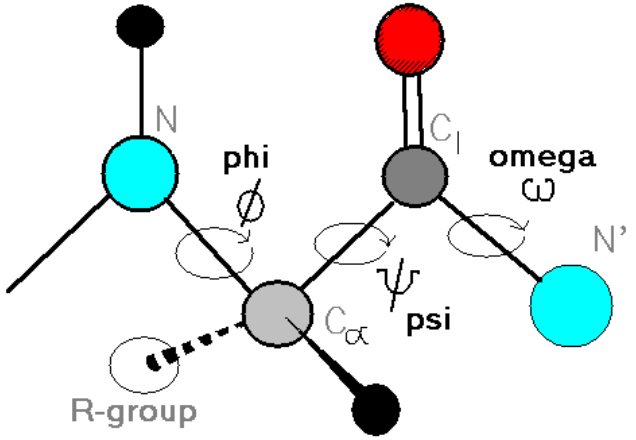
β-strand:

$-180 < \phi < -60$

$180 > \varphi > 60$

α-helix:

φ: ~ -60



External Links

External Links

SECONDARY STRUCTURE

- Secondary Structure Assignments (DSSP)

STRUCTURE CLASSIFICATION

- DALI
- Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists (FATCAT)
- Structural Classification of Proteins (SCOP)
- Protein Structure Classification (CATH)
- Vector Alignment Search Tool (VAST)

STRUCTURE FEATURES

- Computed Atlas of Surface Topography of proteins (CASTp)
- Guassian Network Model (GNM)
- Analysis of interatomic Contacts of Structural Units (CSU)
- Analysis of Ligand-Protein Contacts (LPC)
- SWISS-MODEL : *No external link available*
- Homology derived Secondary Structure of Proteins (HSSP)

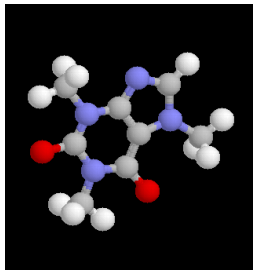
STRUCTURE SUMMARY

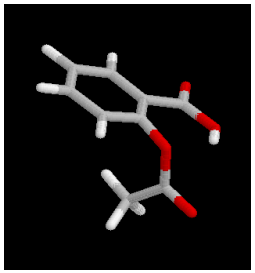
- Molecular Modeling DataBase (NCBI/Entrez) (MMDB)
- OCA Browser (OCA)
- Protein Data Bank Japan (wwPDB Partner) (PDBj)
- Protein Quaternary Structure (PQS)
- Proteopedia
- Biological Magnetic Resonance Data Bank (BMRB)
- Protein Databank in Europe (PDBe)
- PDBsum
- Jena Library

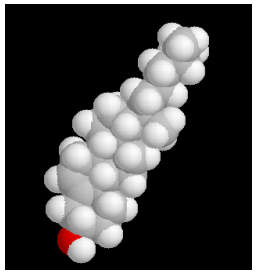
Molecular Graphics

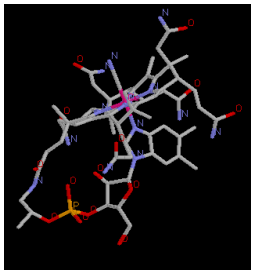
- 是一種可顯現出生物巨分子結構的軟體，包含蛋白質、**DNA**、**RNA**、化學小分子和 金屬等皆可藉由視覺化軟體來呈現他們的結構。

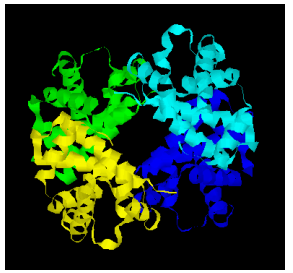
- 視覺化的軟體可以輔助觀察巨分子的結構、作用力、表面特性等。尤其在藥物設計、分子模擬上有很大的應用空間。


咖啡因

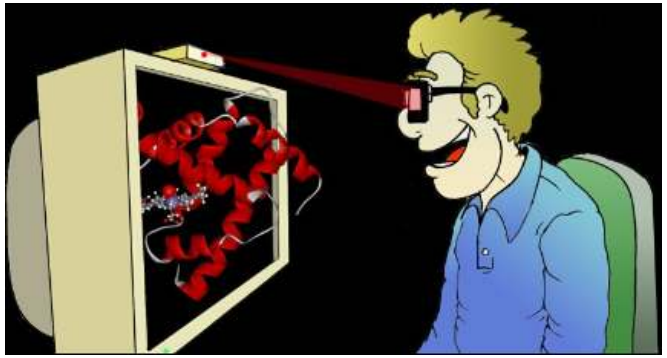

阿斯匹靈


膽固醇


維生素B_{12}


血紅蛋白

Molecular Graphics



- MDL Chime
- RasMol
- PyMOL
- Cn3D
- Swiss PDB viewer
- MOLMOL
- MolScript
- Raster3D
- GRASP
- WebLabViewer

Thank you!